

Technology and the Future of Humanity:

Complete Transcript of the Dialogue

between

Hu Jiaqi, Founder of Humanitas Ark, and Nobel Laureate Michael Levitt

Recently, Mr. Hu Jiaqi, Founder and Chairman of the Humanitas Ark, held an academic dialogue with visiting Nobel Laureate in Chemistry, Professor Michael Levitt, on the theme of "Technology and the Future of Humanity." The two sides had an in-depth exchange of views on a number of topics. In this collision of ideas between top scholars, we witnessed the convergence of science and humanities, reason and compassion. In this era of accelerating technological iteration, humanity's future requires more visionaries to safeguard it. Only by setting aside differences and starting from the common interests of humanity can we achieve the sustainable survival and universal well-being of humankind.

Below is the complete dialogue between Chairman Hu Jiaqi and Professor Levitt. For ease of viewing, the content has been minimally edited, while preserving the original arguments and thought processes of both speakers to the greatest extent possible.

Mr. Hu Jiaqi is the world's earliest pioneer in the systematic study of technological crises and a key architect of its theoretical framework. He has long been committed to raising public awareness of technological risks, calling for rational constraints on technological development, and firmly advocating strict controls on high-risk technologies. Professor Michael Levitt, a Nobel laureate who has made milestone contributions to the scientific field, developed the "multiscale models for complex chemical systems," which successfully bridged the gap between classical physics and quantum physics, turning computer simulations of chemical reactions from concept into reality, and profoundly influencing the development of modern chemistry and life sciences.

When a thinker who maintains a cautious stance toward technology meets a pioneer who has achieved breakthroughs in basic science, what sparks of thought will emerge from this collision of perspectives? Now, let us dive into the full transcript of this top-level dialogue.



I. Exchange on Hu Jiaqi's Academic Theoretical Framework

Hu Jiaqi:

Professor, we would like to divide this into two aspects. First, let me introduce my basic viewpoints and research. Then we can discuss some questions. If you disagree with some of my views, that's fine – we can discuss and exchange ideas.

Today's theme is Technology and the Future of Humanity: A Dialogue with Nobel Laureate Michael Levitt.

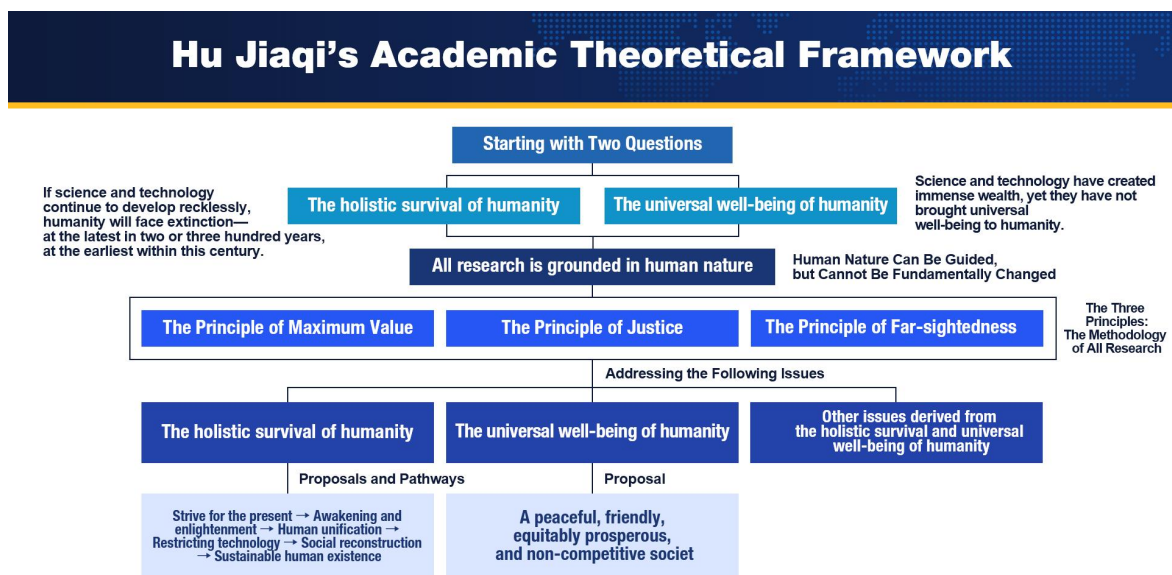
I am here today as the Founder and Chairman of the Humanitas Ark.

I would like to outline the general scope of my research over these many years.

(pointing to the PPT) My academic theoretical framework can be roughly illustrated in this diagram.

I have devoted my life to the study of human issues for nearly half a century — 47 years, to be precise. In the early stage, I spent 28 years completing this 800,000-word book, *Saving Humanity* (pointing to the two volumes of *Saving Humanity*), which was published in China in 2007. After that, I wrote other books, published some articles, wrote letters, and gave speeches, interviews, and produced videos. All my work on human-related issues follows this theoretical framework.

Now let me introduce this theoretical framework.



(pointing to the PPT) My entire research stems from two core questions:

- 1.The overall survival of humanity.
- 2.The universal well-being of humanity.

I believe these are the two greatest issues; there are no issues greater than these.

Regarding the overall survival of humanity, through rigorous and systematic argument,I have reached the following conclusion: if science and technology continue to develop as they are now, they will inevitably lead to

the extinction of humanity – within two to three hundred years at the longest, or within this century at the shortest. I will explain my reasoning later, and we can discuss it.

This means science and technology cannot continue developing in this way. We must solve this problem. Therefore, this is the main topic I want to discuss with you today.

The second issue is the universal well-being of humanity. Science and technology have created immense wealth, but they have not brought universal happiness to humanity. This is not my own statement; it comes from historians.

For instance, L.S. Stavrianos mentions in *A Global History* that human happiness today is even less than that of people in the Paleolithic Age when they first emerged from caves.

Other historians have expressed similar views, drawing on archaeological evidence. Perhaps we can explore this topic in depth another time, but it is not the main focus of our discussion today.

Throughout my research, everything is based on human nature.

I believe that humans are a specific species. We are not gods or immortals; we are simply human beings. Human nature can be guided, but it cannot be fundamentally changed. Any social system designed against human nature will inevitably fail, in my view. We have seen examples of such failures in China's history.

Based on human nature, I have established three major principles: the Principle of Maximum Value, the Principle of Justice, and the Principle of Far-sightedness. These three principles constitute the methodology for all my research. I can explain what these three principles mean later if time permits.

During the research process, we find that if we violate these three principles, our research conclusions might be erroneous. First, we must all agree on the correctness of these three principles. The methodology I have constructed aims to ultimately solve three problems.

The first is the overall survival of humanity.

What is the solution and path to resolving this issue? (pointing to the PPT) As I said, if science and technology continue developing in this way, they will quickly lead to human extinction. How should we solve this? I believe the fundamental solution is to achieve the Great Unification of humanity and strictly restrict the development of science and technology. We should only popularize safe and mature technologies globally.

The implementation path is: (pointing to the PPT)

First, strive in the present: We must take various actions now to avoid technological catastrophes. For example, there is currently the "alignment theory" proposed by Yudkowsky. I think it is useful, but not significantly so — having it is better than having nothing.

Second, promote an awakening movement: Promote a global awakening movement, encouraging humanity to re-evaluate science and technology.

Third, through the awakening movement, achieve the Great Unification of humanity.

Then, strictly restrict the development of science and technology, permanently seal up and even cause to be forgotten high-risk technologies, especially scientific theories.

Rebuild our social systems, restructuring society to achieve the perpetual survival of humanity.

This is the path to solving the problem of 'the overall survival of humanity.'

The second issue is universal well-being.

I have constructed a social framework: to establish a peaceful, friendly, equitably prosperous, and non-competitive society. I believe such a society is most conducive to achieving universal well-being. At the same time, such a society also contributes to the stability of a unified world.

The third issue (pointing to the PPT), we also need to address other issues derived from the overall survival and universal well-being of humanity.

This encompasses all my research, whether it's the 800,000-word book *Saving Humanity*, or other books and articles I have written — all operate within this overarching framework.

That concludes my first point. After hearing this, Professor, do you have any questions you would like to discuss with me?

Levitt:

I understand this very well. I'm not sure that I agree with everything, but I certainly understand what you said. We could go on.

II. "The Three Principles" — An Exchange on Hu Jiaqi's Methodology

Hu Jiaqi:

Alright.

Let me briefly explain my methodology — the three principles.

First is the Principle of Maximum Value.

I believe that the most important values for humanity are, first, survival, and second, well-being. And the overall weight of humanity as a whole is the greatest. Therefore, the holistic survival of humanity overrides all. When the interests of individuals, nations, or countries conflict with the overall survival of humanity, they must unconditionally submit to the latter.

A key slogan of our philosophy is: the holistic survival of humanity overrides all!

You can see this written on the flyleaf of my book: "the holistic survival of humanity overrides all." If someone disagrees with this principle, it is very likely that my subsequent conclusions would be erroneous.

Explanation of The Three Principles



The Principle of Maximum Value

The most important values for humanity are, first, survival, and second, well-being. And the collective survival of humanity carries the greatest weight. Therefore, the holistic survival of humanity overrides all. When the interests of individuals, nations, or countries conflict with the holistic survival of humanity, the latter must be unconditionally prioritized.

The Principle of Justice

Different human groups and individuals, as well as different generations, shall enjoy equal rights to security of existence, the pursuit of well-being, and the allocation of resources.

The Principle of Far-sightedness

- In studying human issues, a sufficiently long-term perspective must be adopted in terms of time, and a sufficiently broad perspective must be taken in terms of space.
- Decision-making must abandon short-term thinking and fully consider the impact of actions on future generations.

Levitt:

I agree with that.

Hu Jiaqi:

Good. Second is the Principle of Justice.

The Principle of Justice means that different human groups and individuals, as well as different generations, should enjoy equal and equitable rights to survival security, the right to pursue well-being, and the right to resource allocation.

This is the Principle of Justice.

If someone believes they must enjoy more than others, that violates the Principle of Justice.

If someone disagrees with the Principle of Justice, many of our research conclusions might also be opposite.

This is my second principle.

Third is the Principle of Far-sightedness.

The Principle of Far-sightedness means that when studying human issues, we must consider sufficiently long time scales and sufficiently broad spatial scales.

Decision-making must abandon short-sighted thinking and fully consider the impact of actions on future generations.

This "sufficiency" goes beyond ordinary imagination. Our human species could potentially survive on Earth for hundreds of millions of years. Dinosaurs survived on Earth for 160 million years, and nautilus have survived for 530 million years. They survived so long because they were strong, and we humans are even stronger than them. If we do not consider issues from such a distant perspective, it is very likely that we, a powerful humanity, will eventually cause our own demise.

In my research, my time scale extends from the beginning of the universe to its end, and my spatial scale extends from beneath our feet to the edge of the universe.

My research covers this vast scope.

When we research with such a scope, it is not to do something at the edge of the universe, nor to do something billions of years away in the future. Rather, we use this scope as a reference point to solve our present issues.

Therefore, the Principle of Farsightedness is not an ethereal, futuristic action plan. It merely uses this perspective as a method to research what we need to do now.

These are the three principles I have established, which form the methodology for all my research.

Professor, you can raise any questions anytime, either now or after I finish speaking.

Levitt:

I understand very well.

III. The Methodology for Arguing that Technology Could Lead to Human Extinction

Hu Jiaqi:

Now, let's focus on the main topic I want to discuss with you. (pointing to the PPT) the issue that many people find difficult to accept: my assertion that the continued development of science and technology will inevitably lead to human extinction soon, within two to three hundred years at the longest, or within this century at the shortest.

How do I argue this? I use a method combining Path Analysis and Defense Limit Testing.

Proving that technology will cause human extinction is difficult; it cannot be falsified, cannot be tested through repetition, and we only have a single sample – our Earth and our one humanity. Therefore, it's hard to use other methods of argument.



**The Argumentation for
Technology Induced Human Extinction:
Pathway Analysis +
Defense Limit Testing**

- Technological development inevitably leads to enhanced destructive power of tools
- Since the Industrial Revolution, technological development has rapidly approached the capacity for extinction
- As Long as Technology Continues to Develop Uncontrollably, the Destructive Power of Means Will Only Grow Greater, Until Extinction-Level Means Emerge
- Extinction means cannot be neutralized (i.e., cannot be defended against)
- The eruption of forces of extinction is unavoidable

Some scientists use probabilistic methods, but I am not very favorable towards them. I find it less scientific.

Let me give a simple example: the overall survival of humanity is incomparably important. Even if the probability is one in ten thousand, multiplied by infinity, it is still approximately infinity — so it cannot be done. Even if the probability is one in a million, multiplied by infinity, it is still approximately infinity. That is why I do not appreciate the approach of calculating the probability of technological extinction of humanity.

Take, for example, Toby Ord from Oxford University. He believes that within one hundred years, the probability of technological extinction is one in six. Hinton, an important contributor to generative artificial intelligence, believes that within forty years, the probability of generative AI spiraling out of control is twenty to thirty percent.

I am skeptical about the scientific validity of his sample. If we were to take his sample seriously, then we must stop artificial intelligence immediately. Because one-sixth is about sixteen percent, and sixteen percent multiplied by infinity is still infinity. Hinton's twenty to thirty percent, multiplied by infinity, is still approximately infinity. This cannot be done. So I believe that method is unscientific. Perhaps I am wrong; this is just my personal view, open for discussion.

Therefore, although technological extinction cannot be falsified, cannot be observed, and the experiment cannot be replicated, extinction risk can still be partially verified and assessed.

Thus, my argument regarding human extinction is approached through two methods combined: extinction pathway analysis and defense limit testing.

How does this argument unfold? My starting point is this: (pointing to the PPT)

First, technological development inevitably leads to an increase in the destructive power of tools.

For instance, in the primitive era, during the Old Stone Age, people used stones and sticks to chase wild animals. These also served as tools for human combat, but their efficiency was very low.

Later came the age of cold weapons, where a sword or a spear could kill one person at a time. Then gunpowder was invented, giving rise to explosive shells, which could injure many people with a single blast.

Later came the mass-energy equivalence equation, which led to nuclear weapons.

Nuclear weapons can destroy an entire city.

So, as science and technology advance, destructive power increases — this is common sense. It needs no proof; it is a historical fact.

I argue this in five steps. This is the first step.

My second step is that since the Industrial Revolution, technological development has rapidly approached the capacity for extinction.

In fact, the acceleration of technological progress began with the Industrial Revolution.

Before the Industrial Revolution, development was slow. The acceleration began in the mid-18th century with the Industrial Revolution and has grown exponentially. In just over two hundred years, humanity has developed science and technology to its current height.

As mentioned, nuclear weapons are already extremely powerful.

Today, we have genetic engineering, nanotechnology, and artificial intelligence.

Many scientists are now concerned that the further development of these technologies — and their destructive potential — is no longer about how many people could be injured or killed. The concern about safety itself is whether they could cause human extinction.

We have reached this point. Although these technologies cannot yet cause extinction, they are approaching that capability. And time is accelerating; the future will only move faster, not slower. This is the second step of my argument.

The third step is that as long as technology continues to develop uncontrollably, the destructive power of these tools will inevitably grow even greater, until eventually extinction-level means emerge.

What is the reasoning behind this?

Even if today's nuclear technology, genetic engineering, nanotechnology, and artificial intelligence cannot lead to extinction, as long as technology continues to advance, something even more powerful will emerge. That something will eventually possess the capability to end humanity.

And after the first extinction-level technology appears, newer ones will follow. These new extinction technologies will increasingly become accessible even to corporations — or even to individual scientists. While the actions of a nation are somewhat controllable, the actions of individuals are very difficult to control. There are always people willing to commit extreme evil.



Irresistible Forces of Extinction

“

Due to time lags and technological bottlenecks, when humanity faces with technologies with destructive power, it is extremely difficult to develop effective countermeasures in time. The widely popular "alignment theory" in the field of AI today suffers from multiple logical flaws.

1. "Alignment theory" can only play a limited role in restraining the dangers of AI. A single instance of misalignment could lead to the extinction of humanity.
2. "Alignment theory" fails to overcome the weaknesses of human nature. The possibility that extremists might use technology to endanger human survival is difficult to preclude. Even if technological alignment could be achieved every time, humanity itself remains incapable of alignment.

The fourth step is that extinction-level means cannot be neutralized — they cannot be defended against.

Some means simply cannot be countered. Take nuclear weapons: once a nuclear explosion occurs, it happens; casualties occur; there is no way to neutralize nuclear energy — no such mechanism exists.

But some means can be countered. For example, if a virus is released causing a pandemic, corresponding drugs can be developed to treat it — this can be counteracted. However, even when neutralization is possible, any countermeasure requires time for development and deployment. During this time lag, humanity is exposed to the threat of extinction.

As extinction-level means multiply, and as the variety of such means increases, and given that humanity's future is extremely long, this time window will be infinitely expanded. Even if neutralization is possible — perhaps even nuclear weapons could eventually be neutralized — the process of research and development still takes time.

Because of this time lag, and because the period without effective countermeasures is prolonged, humanity will be exposed to the threat of existing extinction-level means without solutions for a very long time. Over time, disaster is inevitable.

This is what I call defense limit testing. Assuming all means can be countered, there is still a research process involved.

That is the fourth step.

And what is the fifth step? (pointing to the PPT) The fifth step is that the outbreak of extinction-level forces is inevitable.

Science and technology are inherently uncertain. What we often consider beneficial may turn out to be the most harmful. Moreover, the dark side of human nature means there will always be people seeking to commit extreme evil — every era has such individuals. Once technology reaches the level capable of causing human extinction, extinction forces will inevitably erupt. There are three pathways:

First, extreme means could be maliciously used by evil individuals.

Second, accidents in laboratories could lead to the release of extinction-level forces. Extinction-level technologies are advanced, and advanced technologies carry even greater uncertainty. Scientific research is an exploration of the unknown, and since it is unknown, there are things we do not yet understand. When technology reaches a very high level, something unknown could emerge from that realm — a monster that could annihilate humanity. That is the second pathway.

The third pathway is the unintended misuse of technological products.

Many technologies, such as Freon, were initially considered beneficial, but their use led to ozone depletion. Such situations will become increasingly common.

Through these five steps, I have completed my pathway analysis and defense limit testing.

Pathway analysis — this is the pathway. (pointing to the PPT)

Defense limits (pointing to the PPT) — I assume that all means can be countered. But developing countermeasures takes time, and humanity's future is extremely long, so the window period without effective defense becomes very long. And because the overall survival of humanity is incomparably important, in addressing this issue, it is not just a matter of simple conclusions — it is also a matter of decision-making logic. This (technological development) cannot continue.

That is my argument.

Professor, is there anything you would like to discuss with me?

Shall we discuss first, or shall I continue with the presentation?

Levitt:

Overall, I agree with your views.

For example, technology does carry certain risks, and the future also holds certain risks. But the future is unknown. We know that human extinction in the future could be caused by technological factors. However, it could also be caused by other factors. For instance, natural forces are a very important factor. So the possibility of extinction could also come from nature in the future.



IV. Natural Threats to Human Extinction

Hu Jiaqi:

My argument on human extinction is actually divided into two parts. In this presentation, I am only discussing technological extinction. I have also extensively studied natural threats to human extinction, covering approximately 200,000 words.

I classify natural extinction threats into several levels.

The first level is the cosmic level. Cosmic-level threats are those from which humanity cannot escape, such as the end of the universe. The end of the universe is an unavoidable fate for humanity. It is a foregone conclusion.

Other examples include black hole accretion, supernova explosions, the potential accretion of micro black holes, independent planet impacts, and alien invasions. These are cosmic-level dangers.

My conclusion is that these events are either billions of years away, or their probability is so extremely small that it can be ignored.

That is the cosmic level.

The second level is threats within the solar system.

One of these is unavoidable: in about five billion years, the sun will become a red giant. It will expand dramatically, and humanity cannot escape unless we emigrate beyond the solar system.

Another threat is large asteroid impacts. However, the probability of such impacts is extremely small.

The prevailing view is that a ten-kilometer-diameter asteroid struck the Yucatán Peninsula in Mexico and caused the extinction of the dinosaurs. But I believe that a ten-kilometer asteroid might have caused dinosaur extinction, but it would not have wiped out humans.

Human survival capability is far stronger than that of dinosaurs. We are not defenseless against nature, nor do we simply gather food; we produce it creatively. Thus, ordinary asteroid impacts cannot cause human extinction. An asteroid over one hundred kilometers in diameter might be able to cause extinction—though I still doubt it could.

Such events happen only once every hundreds of millions of years. Within the solar system, there are other factors, such as solar activity — sunspots, magnetic storms — and reversals of Earth's magnetic field, which could leave us vulnerable to cosmic rays for a period of time.

However, threats at the solar system level are either many billions of years away, or extremely low in probability, or simply incapable of causing extinction. So solar system-level threats are not something to be overly concerned about. That is my second level.

The third level is Earth-level dangers.

For example, super-volcanoes and super-earthquakes. In history, entire oceans have been polluted, leading to several mass extinctions. There are also extreme climate fluctuations — times when polar ice melted completely, and times when even the equator was covered in ice. But none of these are sufficient to cause human extinction. They may kill many people, but not all of humanity. Human survival capabilities are too strong. In winter, we enjoy heating indoors; in summer, we use air conditioning. We use technology to cultivate food in any environment — even in the Arctic Circle, food can be grown in greenhouses. So while such events may cause many deaths, they cannot lead to human extinction.

That is within the Earth-level category.

The fourth level includes threats related to evolutionary patterns and biology, such as super-viruses, super-bacteria, and even the possibility of human degeneration.

These could cause many deaths, but they cannot lead to extinction. My definition of human extinction is that not a single individual remains. So my conclusion is that natural threats are either billions of years away, or their probability is so low that they can be entirely ignored.

That is the case with natural threats.

V. Human-Caused Threats to Extinction

Hu Jiaqi:

The second major category is threats originating from humanity itself.

I have also analyzed these extensively. Science and technology are just one factor, but they are the decisive factor. Compared to this factor, all others pale in significance — its weight is enormous. In terms of causing human extinction, the timeline is long at most two to three hundred years, and short within this century.

As Hinton said, the probability of AI spiraling out of control within forty years is twenty to thirty percent. Is that probability high or low? The overall survival of humanity is incomparably important — it is above all else, it is an infinite value. An infinite value multiplied by twenty to thirty percent is still infinite. This is what Hinton, a leading figure in generative AI, said.

As a decision-maker, this is not a scientific question; it is a philosophical question of decision-making. When faced with a twenty to thirty percent chance of human extinction in the next forty years, what decision do you make? Stop or continue?

As I mentioned earlier, Toby Ord, the Oxford philosopher, said that the probability of technological extinction within a hundred years is one in six, about sixteen percent. I do not know how he calculated that. But once this probability is known, do you feel a sense of fear? Sixteen percent is very large — multiplied by infinity, it remains infinite. Can we continue moving forward?

Professor, feel free to interrupt with any questions.

VI. Discussion on Restricting Technological Development

Levitt:

(pointing to the axis drawn by Levitt himself) Two million years ago, human beings emerged. This is the timeline (horizontal axis), from two million years ago to the present. And this axis (vertical axis) represents the advancement of civilization, population growth, and technological development. Some human species have gone extinct: Neanderthals, and others — this is history. We can ask the question: how likely is the extinction of Homo sapiens? In the early days, extinction was easy, because technology was primitive and the population was small. We can also look at how species go extinct. Before agriculture, before animal domestication — this is where technology begins: farming, domestication, and so on. Technology made it easier to survive asteroid impacts and volcanic eruptions. With farming, the population grew, and technology made them harder to wipe out.

Now I would like to ask Mr. Hu a question: If you could go back in time and stop technological development, at what point would you choose to stop it?

Hu Jiaqi:

Alright, let me respond to this question.

In the early days, humans were very vulnerable. Their vulnerable is reflected in two aspects when facing nature. First, the population was too small. Second, their ability to cope with nature was too weak.

When I say humans, I mean Homo sapiens. There is still some debate about when the evolution of Homo sapiens was completed, but it is generally believed to have been completed between 100,000 and 200,000 years ago.

When I wrote this book *Saving Humanity*, I wrote that Homo sapiens completed evolution around 50,000 to 60,000 years ago, because that was the information I had at the time. It has now been twenty years since I finished writing the book. Over these twenty years, there have been many new discoveries and archaeological excavations, pushing the timeline back — The new estimate suggests that Homo sapiens completed evolution 100,000–200,000 years ago.

About 10,000 years ago, humanity entered the agricultural era. Entering the agricultural era was predicated on the domestication of animals and plants.

Archaeologists have discovered a phenomenon: thousands of years before humans settled down and entered the agricultural era, they already had the ability to do so. Their intelligence was sufficient, and they could already domesticate some animals and plants.

So why did they delay entering the agricultural era for several thousand years?

This is a perspective from Stavrianos. The evidence I just mentioned also cites the research conclusions of two famous American archaeologists. Stavrianos's explanation is that although humans were still in the Paleolithic era at that time, they had enough food, and they could enjoy a leisurely lifestyle. He believed that the leisure time of people back then was even greater than that of archaeologists today.

People back then did not worry about food. Their level of food security was even higher than it is today. I experienced the Cultural Revolution, and at that time, people were starving to death.

Stavrianos clearly elaborates on these points in his *A Global History*. And other historians also discussed this issue.

This actually brings me to my second topic. I won't discuss the second topic with you today, because we may not even finish the first one.

As I mentioned at the beginning, my research is approached from two angles. The first issue is the holistic survival of humanity. The second issue is the general well-being of humanity. Science and technology have created infinite wealth, but failed to bring universal happiness to humanity. If there is an opportunity in the future, we can discuss the second issue as well. But today, let us focus only on the first topic.

Now, let me answer your question.

Once humans entered the agricultural era, it became very difficult for natural forces to wipe them out.

The reasons are as follows:

First, the human population became larger.

Second, humans were already able to creatively obtain a great deal of food.

When ordinary natural forces could no longer wipe out humanity, the main threat to humans came from ourselves.

Levitt:

If we stop the technology here, we would increase extinction. This extinction is from earth threats. Also there is extinction from solar threats, as you say, also extinction from cosmic threats. Maybe if we look at the future, we would also have other dangers, and if we stop technology, it will be easier for us to be extinct, rather than difficult to be extinct.

Hu Jiaqi:

Alright, let me respond.

In fact, after humans entered the agricultural era, first, humans could not yet cause their own extinction. The means were too primitive—in the age of cold weapons, one sword or one arrow could only kill one person.

It was impossible to wipe out all of humanity.

Levitt:

I mean, the plague of the 5th century, viruses are killing fifty percent.

We could be extinct, we were very lucky.

Hu Jiaqi:

On this point, I may have a slightly different view from yours.

I believe that plagues are closely related to population density. For example, during the 14th century, the Black Death in Europe killed a large number of people. But the Jews survived the plague. Why? Because the Jews lived in their own separate communities having very little interaction with others. They could maintain hygiene and manage their environment well. So the Jews survived the Black Death, while other Europeans did not.

Whether it's a plague, a bacterium, or a virus, its spread is closely related to population density. A naturally occurring plague, at its worst, might kill 99% of humanity—but the remaining 1% could still survive.

Because I have a definition: when I talk about extinction, I mean: not a single person remains—there is no possibility for humanity to come back. In my book, I define extinction as no individuals left, just like the dinosaurs—When there is not a single individual left, this species has no future.

For example, during a certain period in China—the Eastern Jin and Northern and Southern Dynasties—how many people died in northern China? Seventy to eighty percent died, leaving only twenty to thirty percent. Most of that was due to slaughter.

Do you know the term "two-legged sheep"? When there was no military rations, those ten thousand people became your rations. The common people were eaten. And there was even a preference for taste—young children tasted best, then women, then men, and finally the elderly—the elderly were the least tasty.

China has two such periods in its history: One is the Sixteen Kingdoms period of the Eastern Jin. The other is the Five Dynasties and Ten Kingdoms period. These two periods are the most brutal in Chinese history. There was literal cannibalism. When there was no military food, people were allocated as rations.

Levitt:

What year was it?

Hu Jiaqi:

It was before the Song Dynasty — the Five Dynasties and Ten Kingdoms period.

The Five Dynasties and Ten Kingdoms occurred between the Tang and Song dynasties.

But that was not the worst. The worst was the northern region during the Eastern Jin and Sixteen Kingdoms period. That's how miserable northern China was.

Levitt:

What year was it?

Hu Jiaqi:

Between 300 AD and 430 AD. This was the darkest period in Chinese history. In less than a hundred years, the population dropped by seventy to eighty percent.

Pandemics stop when population density decreases, because the transmission routes are cut off. Just like with COVID-19 a few years ago — why did we need quarantine? It was to stop people from moving around.

There are three possible ways a pandemic can stop: First, herd immunity. Second, when the density becomes low enough, it stops on its own. Third, the development of effective treatments.

So a pandemic cannot go on forever — unless it is specifically designed to do so — it cannot kill humans indefinitely. When the transmission density becomes very low, it loses its ability to kill. That's the nature of pandemics.



VII. The Optimal Time to Limit Technological Development

Levitt:

(pointing to the axis drawn by Levitt himself) If we stop here, if civilization stop here (referring to the Stone Age). It is bad.

Hu Jiaqi:

Actually, I would not advocate stopping civilization at this point. If I were a person living at that time, I might feel happy. But having come this far, I don't want to go back. Why? Because people back then lived a very hard life. Although they were better off than other animals by comparison, their standard of living is incomparable to that of modern humans.

Levitt:

What will happen later, we actually don't know.

Hu Jiaqi:

This is what I was just saying — we need to reason it through.

The claim that technology could end humanity is unfalsifiable.

Levitt:

I agree. But what also would be possible is that there is a non-human technology that will destroy humankind if we don't develop our own technology. Have you read the book Three Body? Three Body, the science fiction. The

network through the threat from another civilization , They need to develop technology to fight against the threat. If we stop technology, what would happen? I don't know.

Hu Jiaqi:

I have read the book: *The Three-Body Problem*. I've watched the TV series too. It was written by the Chinese author Liu Cixin. It has a strong sci-fi sensibility. But I think I am doing scientific research, not science fiction.

Levitt:

You've also talked a lot about happiness.

Hu Jiaqi:

Professor, let me finish what I wanted to say. Let me complete my argument. I believe it would have been better if humanity had stopped two hundred years ago. More or less — two hundred years ago would have been quite good.

By then, humans had solved the problem of food.

For a long time, even after entering agricultural civilization, we hadn't truly solved food security. It was only after the Industrial Revolution that productivity exploded. That productivity was tied to relatively simple tools — tools that were intuitive. The more hidden technologies are, the easier they can harm humanity. The less hidden, the more intuitive — and then easier to use for good.

So I think fifty years after the Industrial Revolution would have been enough — That would have been the optimal period. By then, we had solved food and clothing. Look at the spinning machine — Before that, weaving cloth was very slow, productivity very low. Rice planting was done by hand. Transport was primitive — horse-drawn carriages.

I believe the key turning point on which the potential for human extinction to appear is when Einstein's mass-energy equation was proposed. That's when people realized just how enormous the energy contained in science and technology could be. But we can't take 1905 (referring to the year he introduced the mass-energy equivalence equation) as the cutoff point for restricting technological development we would need to push it back. There should be a safety period.

There's no perfect 100% safe period, but I would prefer to set the cutoff at fifty years after the Industrial Revolution. By then, people had solved food and clothing. It was also very safe. Compared to earlier times, there was happiness. Happiness is born from comparison — that is my conclusion.

Levitt:

In that time that you would like to stop technology, there are still many many children are died, many women are treated very badly, It is the time when many people are very unhappy

Hu Jiaqi:

I agree. First, regarding the many children dying from illness — that just means technology hadn't reached a high enough level. If humanity wants to survive forever, certain sacrifices must be made. Without sacrifices, it's impossible to ensure humanity's eternal survival. That's my first point.

Second, regarding women not being treated fairly — that is a matter of social systems. Social system issues need to be discussed using a different kind of logic. In fact, even today, gender inequality still exists around the world. That problem has not been fully resolved — just like Black people in America.

By the 1960s, it had already been a hundred years since Lincoln called for the emancipation of black slaves. Yet Martin Luther King Jr. was still there, delivering his "I Have a Dream" speech.

So these things take time — that's a separate issue. If we want to discuss that, it would be a second discussion. This is the first discussion we are having.

VIII. How to Limit Technological Development

Levitt:

Let's discuss something different. Let's say if we like to stop technology development. It seems to me that it is impossible to do that.

Hu Jiaqi:

This is quite difficult, but I have my own approach.

First, let me make a point: the expansion of human political entities is a major trend.

In ancient times, there were only tens of thousands of humans. China's first dynasty was the Xia Dynasty, which covered only a small part of today's southern Shanxi and northwestern Henan. The Shang Dynasty was a bit larger,

and the Zhou Dynasty larger still. Then Qin Shi Huang unified China, making it even larger. The same pattern happened around the world.

Look at China today — 1.4 billion people, such a vast territory. Russia is even larger. And the former British Empire, right? The expansion of the British Empire, the expansion of the Mongols, right? At one point, the British Empire seemed to own the whole world. Of course, it later stopped, for various reasons.

What I mean is, the expansion of human political entities is a major trend. It has ups and downs — times when it shrinks and times when it grows. Right now, it's in a period of decline. But I think it will soon increase again. Russia (the Soviet Union) collapsed — the Soviet Union was once the largest. That's true. But overall, the trend is toward expansion.

And there is only one reason that prevents the expansion of human political entities. What is it? Only one material condition: inconvenient transportation and communication, which makes it impossible to govern a large region. That is the only material condition that prevents expansion. There is no second condition.

The current stagnation has to do with the mutual deterrence of nuclear weapons, and with the fact that people have not yet reached a consensus.

Once a consensus is formed, human unification is possible. There are too many examples in history. For instance, East and West Germany naturally reunified. The United States originally had thirteen states, and those thirteen states voluntarily united. There have been cases of voluntary unification.

So, the key is to clearly explain to all of humanity the rationale that Great Unification must be achieved.

Levitt:

You mentioned that human groups are getting larger and larger — I agree with that. But if we want to stop technological development, we need an entity to take on this role. For example, look at two large political entities like China and the United States — how many years will it take for them to achieve unification? Fifty years, a hundred years, three hundred years — we don't know. But only after such a unified body is formed can we possibly restrict technological development; otherwise, the current technological competition is actually very fierce.

Hu Jiaqi:

I think you've hit on a core issue. Right now, I cannot answer this question precisely. But there is something I can do — something that would benefit this process, that would contribute to its formation.

What is it? We need an awakening movement.

In human history, (right now) we need an awakening movement. We need to explain this principle clearly to everyone. Once people have a general understanding of this issue, multiple unifying forces will emerge.

Let me give you an example. I can casually design several scenarios that I think could help bring about the unification of humanity. Suppose that today, at this very moment, China and the United States firmly embrace this principle — then the two countries together would have a sixty to seventy percent chance of making it happen (referring to the Great Unification of humanity). If Russia is added, the chance rises to seventy to eighty percent. If India is also added, the chance reaches eighty to ninety percent.

Just imagine: China, the United States, Russia, India. India has the largest population — even larger than China. If we convince these four countries, there is a very high probability of success. Not guaranteed, but very likely. And what's more, these four countries are geographically connected.

If all four agree, who would dare disagree? Small countries would be the most willing, because small countries are always being bullied. Just like when the League of Nations was established — everyone had high hopes for it.

Which countries welcomed the United Nations? The five permanent members of the Security Council welcomed it, because they had special privileges. But the League of Nations was different. The League of Nations had no veto system, so it was the small countries that welcomed it most. Because small countries are constantly bullied by large ones, and they don't want to remain in such an environment. So as long as this awakening movement is done well, promoting it is not difficult.

Historically — and you, Professor, know Western history much better than I do, because I am from the East — there was the Renaissance period, there was the Enlightenment period. Without the Renaissance, there could have been no Reformation by Martin Luther. Without the Enlightenment, there could have been no United States of America — the first political regime in human history with separation of powers, right?

So we need an awakening, an enlightenment. Now, our Humanitas Ark has been working on this all along. I have been doing this since I was seventeen years old. I have written letters to so many people, recorded so many speeches, and even tried to become the richest man — putting all my money into this — just to do this one thing.

IX. Discussion on Artificial Intelligence

Levitt:

Let me ask you a different question. We are now on the era of AI, becoming more and more powerful. Will the power of AI make world unification easy, or more difficult?

Hu Jiaqi:

In fact, AI is like a sharp sword hanging over our heads. It could fall at any moment. AI is developing too fast. Yudkowsky came up with the alignment theory, and many people have praised and recognized it. But I am quite unimpressed with this alignment theory.

Right now, alignment theory is the hottest topic. My view is this: this product might be alignable, that product might not be alignable. Even if all machines could be aligned, humans themselves are not aligned with each other. Many scientists deliberately do bad things. With the tools available, it becomes even easier to do bad things intentionally.

There are plenty of people who want to destroy all of humanity. They just don't have the capability yet. When a scientist has that capability, it becomes very easy for them to do it (to extinguish humanity) if they want to.

So, does alignment theory have any use? It might help a little, but it cannot help much. So originally I thought Nick was my bosom friend, but now I no longer consider him as such.

I can show you this slide about Nick Bostrom right here. I started researching this in 1979. I kept at it until 2007. The results of my research are these two books. After they were fully published, I sent them out to the world.

Once they went out to the world, two extremes emerged. On one hand, some people supported me. On the other hand, the vast majority opposed me and even banned my books. Back then, when I said that technology could destroy humanity, people said I was worrying without cause.

And then which year did things change? In 2013 — six years later — I suddenly came across a report. This report was led by Nick Bostrom of the Future of Humanity Institute at Oxford University. He said something in that report — an official report, involving many scientists and philosophers

He said that humanity could go extinct as early as the next century, and that the main culprit for human extinction is science and technology. That struck a chord with me.

The report gathered many scientists, philosophers, and entrepreneurs for discussion. They said that humanity could be extinct as early as the next century, and the chief culprit for human extinction is science and technology. I was overjoyed when I saw that. I wrote him a letter and also sent him a book, but he never replied to my letter. The book should have arrived, because later when his staff came to Beijing, they specifically visited me. It's just that he never wrote me back. And I published an article in a magazine and online, titled "Finally I've got a Bosom Friend." I thought he was my bosom friend.

But after all these years, looking back today, I realize he is not. Why? First, what he said today is different from his previous conclusion. He said that technology causing human extinction is highly uncertain. With

such high uncertainty, you don't need to worry about it. Second, he has introduced the alignment theory into his research. I think the alignment theory is just a small trick. Having this trick is not a bad thing, but it cannot achieve true alignment — it will never be fully aligned.

What is his version of alignment about? I don't know if the professor is familiar with this alignment theory. What does this alignment theory mean? It means that after artificial intelligence evolves from today's weak AI to strong AI, AI will gain human-like thinking abilities, and then two types of mistakes could occur.

The first is doing harm with good intentions. For example, you give it an instruction, and it follows your instruction, but ends up doing bad things with good intentions. Let me give you an example: It might think that the environment needs to be made perfectly clean, but since humans are the number one destroyer of the environment, it might decide to wipe out humanity. So that's the first — doing bad with good intentions.

The second concern is: What does it worry about? It worries that after AI develops to a certain point, it becomes incredibly smart. It looks at humans and thinks, "How can you be so stupid?" "What I can compute in one second would take you a thousand years." Then it will look down on humans. Higher species despise lower species, and even treat lower species as food — that is an iron law of nature. So it will rebel.

Levitt:

Let me ask you a question: imagine this is a special button, and if you press it, all AI is frozen now. Would you press it?

Hu Jiaqi:

I would press it, but such a thing cannot exist — it will never happen.

Levitt:

If it is possible, would you press it?

Hu Jiaqi:

I would press it. Because even without it, our lives are already good enough. In fact, many people long for a life like the Peach Garden.

X. Discussion about the Future World

Levitt:

I think the reason people yearning for life in the Peach Park is that they don't understand life in the park.

There's terrible disease, suffering, slaves, and terrible torture. I think everyone thinks that it was better 20 years ago or 100 years ago or 1000 years ago. That is because on mind they can't imagine. They think of all the good things. If you would press the magic button, and maybe AI will save humanity.

Maybe AI has the ability to bring us Peach Garden. Maybe AI has the ability to unite the world. Maybe because of AI, all the countries will become one. I don't know. I'm just saying there is uncertainty. We cannot be sure that stop something now is right. Firstly, it is impossible. I agree we cannot stop. But we cannot be sure that stopping is the right thing.



Hu Jiaqi:

Alright, let me answer this question.

We have come from primitive times to today. We have seen many good things, but we have also seen today's problems. Yet today we are faced with a problem that must be solved: humanity may soon be driven to extinction by science and technology. This is too big an issue. Faced with such a huge problem, we must solve it. Only by restricting science and technology to a considerable degree can this problem be solved. So the Peach Garden for us today is not a choice we make willingly — it is a forced choice. We have no choice but to do this.

Of course, the future Peach Garden will be different from the Peach Garden of the past. We must incorporate into that Peach Garden the safest and best things from all of human history, not just oxen plowing fields and horses pulling carts — not that kind of Peach Garden. That is my design for the future.

Since this is not the main topic today, I had not planned to discuss it at length. But since we have come to this point, let me say a few more words. Our future society should be a peaceful, friendly, equitably prosperous, and non-competitive society, where existing safe and mature technologies are widely disseminated across the globe. I believe that widely disseminating existing safe and mature technologies around the world would be more than enough to ensure that all of humanity enjoys prosperity and well-being.

This has been validated on the land of China. China is a large country with over 1.4 billion people, yet its arable land area is smaller than that of the United States and even smaller than that of India. India has only one-third the land area of China, but its arable land is larger than China's. Yet today, China is able to achieve food and clothing security because it has some fairly good agricultural technologies — and these agricultural technologies should be preserved. Also, technologies that pose no harm to the holistic survival of humanity — such as tractor plowing and combine harvesting — can also be preserved.

The future Peach Garden is different from the Peach Garden of the past. The Peach Garden of the past was one that had not experienced the Industrial Revolution. But the Peach Garden of today, having experienced the Industrial Revolution, can discard the bad and keep the good. So it will be more than enough to ensure the

prosperity and well-being of humanity. It is a garden we design ourselves — not a forced choice.

XI. Discussion on the Issue of Happiness

Levitt:

I'm interested a lot in what makes happiness. And after thinking about what makes happiness for a long time, I realize that for happiness, you must have happiness, unhappiness, happiness, unhappiness. Because if you only have happiness, it isn't good. So you need to have unhappiness, happiness, unhappiness, happiness, unhappiness, happiness. In this way, you can have a happy life. This is what I think. What do you think?

Hu Jiaqi:

I have my own definition of happiness. I particularly agree with what you just said: sometimes we are happy, sometimes we are not.

I believe happiness is attained through pursuit and generated through comparison. If a person has no goals to pursue, it is difficult for them to experience happiness. Without comparison, they also lack a sense of well-being. For example, children in wealthy families who eat delicacies daily may lose their appetite, while a poor farmer who occasionally enjoys meat feels it is incredibly delicious.

This is a form of longitudinal comparison. There is also horizontal comparison: For example, when I see others driving luxury cars and living in mansions, while I ride a bicycle to work and live in a small dormitory, my sense of happiness diminishes. Therefore, happiness first arises through pursuit.



Second, it emerges through comparison.

Moreover, the goals of pursuit vary at different life stages. For instance, during school, my goal was to achieve good grades and gain admission to a prestigious university. After graduation, I hoped to secure a good job. At my current stage, I need to support my children, and I wish for them to succeed. If my child were admitted to a renowned university, I would feel deeply fulfilled.

Happiness arises through pursuit. At every moment, people have multiple goals. When does a person feel a sense of happiness in the pursuit of something?

And when does the pursuit of other things bring pleasure rather than true happiness? When they make progress toward their primary goals, psychological satisfaction emerges, creating a sense of happiness.

For example, if I recently lost my job, I would feel terrible and lack happiness. However, if a friend treated me to a delicious meal and offered comfort during this time, I might feel momentarily joyful and happy, but this is not true happiness—it is merely pleasure.

So there are two points:

First, happiness is achieved through pursuit. But what do we pursue? It is achieved when you gain something while pursuing your main goal at that stage.

Second, happiness is achieved through comparison — it can come from comparing with the past, with future expectations, or with the people around you.

Why did Stavrianos say that people today are less happy than people in the Paleolithic era? Because when ancient people compared, they felt very happy. Why? First, everything they saw was lower-level animals — even tigers and lions couldn't beat humans.

Because humans could just set a trap and catch them. So when they compared themselves to all the animals, they felt a strong sense of happiness. Second, they had a stable food source and plenty of leisure time. They didn't have to constantly struggle to obtain food. They felt that going out to hunt one or two goats a day could feed the family for several days — that gave them happiness. They didn't know about today's world.

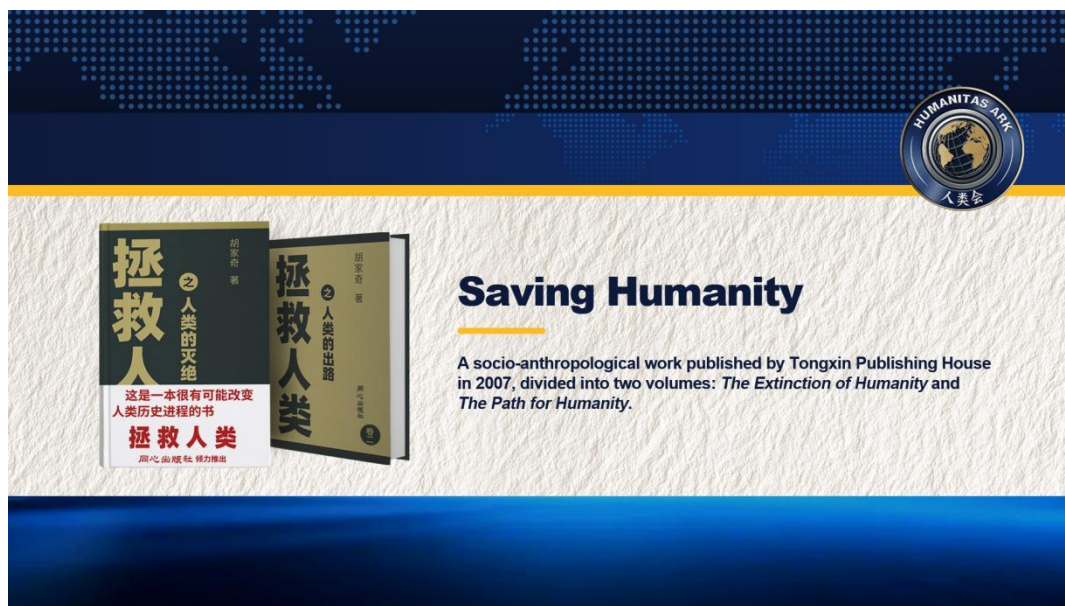
If they had, that happiness would have been gone. So happiness is fluctuating. Happiness is gained when you achieve your main goal at different stages of life. For example, when people grow old, they mostly hope for filial piety from their children. When young, even if conditions are good, they may not feel happy. Also, happiness is achieved through comparison. This is my understanding of happiness.

So what I envision is that (pointing to the PPT), for a future society to achieve happiness, it should be a peaceful one — people should have a more peaceful state of mind. Don't be anxious. If you see someone eating a piece of meat while you only get the bone, and you get upset — you're just being hard on yourself. We should live in friendship, be kind to neighbors, and avoid conflict. Human beings have an eternal tendency toward conflict. Equitable prosperity. Why must we have equitable prosperity? A society with a huge wealth gap — it is inevitably one where a small minority is wealthy while the majority is poor. In such a society, only the minority are happy, and most people are unhappy. That's where resentment toward the rich

comes from. So we need an equitably prosperous society, (pointing to the PPT) as well as a non-competitive society.

Today, science and technology produce such a massive amount of knowledge —every day, such a huge volume of knowledge forces you to learn and to compete. A couple together cannot even afford to raise one child — it's all due to constant comparison.

Therefore, I believe that a peaceful, friendly, equitably prosperous, and non-competitive society is the kind of society that most enables people to achieve happiness, and it is also the most stable society. So I believe that a society of Great Unification is one that can prevent human extinction, and it is precisely also a society that can achieve universal well-being for humanity.



XII. Hu Jiaqi's Struggle Journey for Humanity's Cause

Hu Jiaqi:

Let me tell you the story of what I have done over the years. Why did I decide to dedicate my entire life to this cause?

I come from a rural background. When I entered university, I didn't even know about Einstein. Because at that time, it was the Cultural Revolution, and we had very little access to knowledge. After I entered university, I was exposed to many cutting-edge scientific theories. These cutting-edge technologies made me realize how immensely powerful science and technology are. Under these circumstances, I felt there was a question worth studying: Is it possible for the development of science and technology to drive humanity to extinction? When will it happen? Can humanity be saved? So I set a lifelong goal for myself.

But after graduation, I was assigned a very good job — first at a research institute, then at a government agency. Other people were assigned to mines or factories. But I was determined to work on this cause. I kept at it. I'm very fortunate that our workplace is right next to the National Library. Whenever there was a holiday, I rushed to the National Library. But I still felt time was not enough, so I decided to "go into business".

At that time, I thought to myself: once I earned 500,000 yuan, I would stop. But in my first month, I made over 1 million yuan. A year later, I earned more than 10 million yuan. Back then, 10 million yuan was worth about what 10 million US dollars is worth today. In that first year alone, I earned that much. When I first went into business, my monthly salary was only 186 yuan. That meant I earned in one year what would have taken me thousands of years at my old job. So I felt very lucky and believed I had some talent for business.

By then, I handed over the business to others—my colleagues, who still run it today. I completely withdrew from business operations and focused on writing at home. For six years, I was entirely devoted to writing. I wrote for ten years in total.

When I finished the book *Saving Humanity*, I wanted to promote my work well, just like Nick Bostrom did. I reached out to many media outlets for promotion. But when the media started promoting my book, it was banned from sale. This happened in 2007.

What could I do then? I began writing letters—to national leaders, scientists—and sent them copies of my book. Over the years, I sent out 12 batches of letters, totaling about 1 million. Around 800,000 were emails, 200,000 were paper letters, and I mailed out 10,000 copies of my book.

By 2008, I built my personal website, which is searchable online and available in both Chinese and English. I thought that simply posting things online wasn't enough, so I created a dedicated site for promotion. Later, as the Humanitas Ark (originally the "Save Human Action Organization") formed, we launched another website. Now we have two dedicated websites: my personal one in two languages, and the organization's site in over a dozen languages, promoting globally.

I also published books in various places: in the U.S., Russia, Singapore, and Hong Kong. But the number of people who recognized my work remained small. The promotion results have never been ideal. What should be done?

So I decided: I would become the world's richest person. If I became the world's richest, my voice would carry weight—like Elon Musk, who speaks with authority as the world's richest person. I pooled all my money and launched several internet projects: Yi Lu Lao, Yi Ba Yi Ba Lao, and Zhi Jie Lao. The press conference of my first website was held at the Great Hall of the People. This press conference cost over four million RMB. In the end, I worked on the internet for nearly ten years. It was too difficult because of the first-mover advantage problem. I had burned through most of my money, so I couldn't continue. What could I do if I had no capital left for anything else?

Then I thought: since this couldn't be done, since this failed, I would start an organization and use collective strength to push things forward. This organization already has over 13 million supporters (and has now surpassed 14 million). This is what I have done.

Today's event is also a new approach. Perhaps next month, I will invite another Nobel laureate. Top scientists from around the world need to reach a consensus — that's what I think.



XIII. Convergence of Key Ideas

Hu Jiaqi:

It has been 47 years since I started my research on human issues in 1979. During this time, there have been some points that align with my views. As I just mentioned, my book *Saving Humanity* was published in 2007. In my book, I proposed that extraterrestrials should not be contacted without caution. Two years later, Stephen Hawking raised the same issue. I even wrote a special article about this, titled “The World Is Crazy,” which was published in the Hong Kong magazine *Frontline*.

In 2013, Nick Bostrom issued an official report stating that humanity could go extinct as early as the next century, and the culprit would be science and technology. I wrote an article specifically about this, titled “Finally I’ve Got a Bosom Friend.” I also wrote him a letter, sent it along with my book — but he never replied. However, when his staff came to Beijing, they did visit me.

In 2014, Hawking said that AI could spiral out of control and cause human extinction. Hawking said this in December 2014 — which was seven years after I made that point.

In 2017, Hawking said in an interview with *The Times* that we might need to establish a world government for regulation. That world government is exactly what we just called the Great Unification of humanity. He was a full ten years behind me on this view.

Of course, today more and more people recognize that science and technology could lead to human extinction. But there are still two points that very few people agree with me on. What are those two points? First, science and technology must be strictly restricted. Disseminating existing safe and mature technologies across the globe would be more than enough to ensure that all of humanity enjoys prosperity and well-being. Second, humanity must achieve Great Unification and use the power of a world regime to control science and technology.